Integration of Inertial Information with Vision towards Robot Autonomy

Jorge Lobo
Institute of Systems and Robotics
University of Coimbra
3030 Coimbra, Portugal

Jorge Dias
Institute of Systems and Robotics
University of Coimbra
3030 Coimbra, Portugal

Abstract - **Reconstructing 3D data from images acquired by cameras is a difficult task. The problem becomes harder if the goal is to recover the dynamics of the 3D world from the image flow. However, it is known that humans integrate and combine the information from different sensorial systems to perceive the world. For example, the human vision system has close links with the vestibular system to perform every-day tasks. A computational approach for sensorial data integration, inertial and vision, is presented for a mobile robot equipped with an active vision system and inertial sensors. The inertial information is a different sensorial modality and, in this article, we explain our initial steps to combine this information with other sensorial systems, namely vision. Some of the benefits of using inertial information for navigation and dynamic visual processing are described in the article.**

During the development of these studies a low-cost inertial system prototype was developed. A brief description of low-cost inertial sensors and their integration in an inertial system prototype is also described. The set of sensors used in the prototype include three piezoelectric vibrating gyroscopes, a tri-axial capacitive accelerometer and a dual axis clinometer.

As a first approach the clinometer is used to track camera's pan and tilt, relative to a plane normal to the gravity vector and parallel to the ground floor. This provides the orientation data that, combined with a process of visual fixation, enables the identification of the ground plane or others parallel to it. An algorithm that segments the image, identifying the floor along which the vehicle can move is thus obtained.

I. INTRODUCTION

In human and other animals the vestibular system gives inertial information essential for navigation, orientation or equilibrium of the body. In humans this sensorial system is located in the inner ear and it is crucial for several visual tasks and head stabilization. The human vestibular system appears as a different sensorial modality, that cooperates with other sensorial systems and gives essential information for everyday tasks.

One example of cooperation is between the vestibular sensorial system and the visual system. It is well known that the inertial information plays an important role in some eye and head movements [1]. The eye movements can be classified into two broad classes: movements related with gaze-changing behaviors and movements related with gaze-holding behaviors. The gaze-holding movements try to maintain the image stationary on the retina despite the ego-motion or movements of the head. The information provided by the vestibular system is used during the execution of these movements, as described by Carpenter [1]. However the inertial information is also important for head- stabilization behaviors, including the control of posture and equilibrium of the body.

Inertial information can be obtained through the analysis and processing of the measurements given by artificial sensors. These artificial inertial sensors have some interesting properties since they are non radiating, non jammable and can be assembled in self-contained inertial systems, providing dynamic information through direct measurements.

There are two types of low-cost inertial sensors, both providing a *rate* measurement: the *accelerometers* and the angular rate sensors or *gyroscopes*. The accelerometers provide linear velocity rate or acceleration. The gyrometers or gyroscopes provide measurements on angular rate information or angular velocity. The experiments related with the properties of the vestibular system on primates indicate that their inertial sensors have resolutions similar to the actual low-cost artificial sensors now available - see Carpenter for measurements on the human's vestibular system [1].

Inertial systems have been used on vehicles such as airplanes and naval ships. Self-contained *inertial navigation* uses the double integration of the acceleration sensed in an inertial space. The mechanization of this procedure normally includes some additional constraints on the process. These constraints are related with universal physical laws that must be properly used. Depending on the approach and application, these laws involve the existence and constancy of gravity, the constancy of the speed of light and the accurate measurement of time - see Russel for details [2].

The inertial information can also be useful on applications with autonomous systems and artificial vision. In the case of active vision systems, the inertial information gives a second modality of sensing that gives useful information for image stabilization, control of pursuit movements, or ego-motion determination when the active vision system is used with a mobile platform. This kind of sensorial information is also crucial for the development of tasks with artificial autonomous systems where the notion of *horizontal* or *vertical* is important, see Viéville for one example [3].

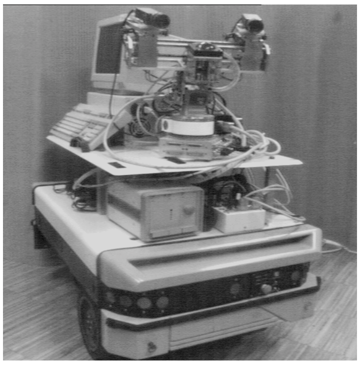

Fig. 1. The mobile system with the active vision system.

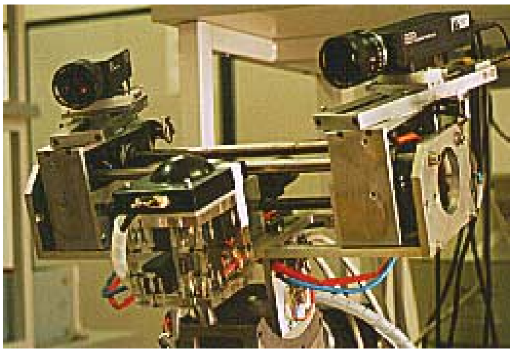

Fig. 2. Detail of the inertial system prototype mounted on active vision head

The inertial system prototype described in this article is used in a mobile robot with an active vision system - see Fig. 1 and Fig. 2 - and unlike some systems that rely on outdoor environments to obtain additional information, like GPS NavSat [4], or on a knowledge of the vehicle's moving space [5], we intend to use the system either indoors or outdoors and in unknown environments.

The following sections describe the mobile system used and a first approach of inertial and vision data fusion, namely in identifying the ground plane.

II. INERTIAL SENSORS

There are inertial systems currently available with high-quality. These systems are normally for aerospace applications and they are expensive when compared with other sensorial systems used in robotics applications. This is one of the reasons why these systems are not so common in robotics applications. However the recent development in solid-state accelerometers and gyrometers has lead to new low-cost sensors. The needs of the automotive industry increased their availability, lowering the cost. These sensors, though not suitable for high performance inertial systems, can provide an inertial system suitable for some robotic and land vehicles applications, at a much more competitive price than their aerospace counterparts.

The new commercially available accelerometers are typically silicon micro-machined sensors and incorporate amplification, signal conditioning and temperature compensation (see [6], [7] or [8] for examples). There are presently three main types of accelerometers which offer reasonably small size. These are the capacitive, piezoelectric and piezoresistive types. The piezoelectric sensors have no DC response, making them unsuitable for inertial navigation systems. In the piezoresistive sensors the acceleration causes a mass to move with respect to a frame, creating stress in piezoresistors diffused onto the supporting flexures, changing their resistance. The capacitive sensors rely on the displacement of capacitive plates due to the acceleration, creating a mismatch in the capacitive coupling. This change is used to generate a signal proportional to the acceleration experienced by the sensor. The capacitive sensors also allow a force balancing servo loop, that keeps the sensor at its 0g position, increasing sensor linearity. These devices are built so as to have a sensing axis and reduced off-axis sensitivity. Some are tri-axial, incorporating three accelerometers in one precision aligned ready-made sensor. These sensors have different measurement ranges from ± 1g up to ± 100g.

The traditional gyroscope involves precision mechanical parts and, although more accurate, is very expensive and bulky. Optical gyroscopes such as the ring laser and the fiber optic gyro measure angular rate of rotation by sensing the resulting difference in the transit times for laser light waves traveling around a closed path in opposite directions [9]. Though more accurate, these are still rather expensive when compared to the lower-cost gyroscopes that measure the angular velocity based on the measurement of the Coriolis force when a rigid body is submitted to angular velocity. The gyroscope designed by Murata [10] uses an equilateral triangle prism made from elinvar elastic invariable metal, which is fixed at two points. Three piezoelectric ceramic elements are attached to the faces of the prism, one on each side. The prism is forced to vibrate at about 7 kHz by one of the piezoelectric elements, and the other two are used for measurement. When there is no rotation all piezoelectric elements detect equally large signals. When the prism is turned it gets twisted, so the detectors receive different signals. The difference between signals is processed by the internal electronics to provide an output voltage proportional to angular velocity. There are other devices based on the same principle, but with somewhat different geometry, such as the vibrating cylinder and the tuning fork designs.

A basic inertial system can be assembled with the combination of these type of *accelerometers* and *gyroscopes*. Using three accelerometers (or a tri-axial sensor) an acceleration vector is defined. However this vector must be referenced to a coordinate system which can be maintained or defined in precise manner. The nature of this reference system depends on the nature of the vehicle. For land vehicles or active vision systems the reference can be an Earth referenced coordinate system. Mechanical inertial systems rely on a gyro stabilized platform to keep the accelerometers aligned to a known reference. In strap-down systems the gyroscope enables the tracking of rotations done in the inertial space, providing the necessary reference for the acceleration vector [9].

A theoretical system based in these principles and using three accelerometers and three gyroscopes was presented by Viéville in [3]. In that article, methods of extracting the motion and orientation of the system from inertial information are derived theoretically but not directly implemented in a real system. Barshan in [11], describes a vibrating beam gyroscope, used for short range orientation estimation for outdoor vehicles, enabling wider spacing of navigation beacons, and later in [12] describes an inertial system for mobile robots and some field tests.

III. A PROTOTYPE OF AN INERTIAL SYSTEM

To study the integration of the inertial information in artificial autonomous systems that include active vision systems we decided to develop an inertial system prototype composed of low-cost inertial sensors. Their mechanical mounting and the necessary electronics for processing were also designed. The sensors used in the prototype system include a tri-axial accelerometer, three gyroscopes and a dual axis clinometer.

A. The sensors

The tri-axial accelerometer chosen for the system, while minimizing eventual alignment problems, didn't add much to the equivalent cost of three single-axis sensors. The device used, 34103A model from Summit Instruments, is a capacitive accelerometer with a range of ±2,5g, that employs three polysilicon surface micro-machined sensors. The capacitive sensing is enhanced by a force balancing servo loop that minimizes errors [7].

In order to keep track of rotation on the x, y and z axis three gyrometers were used. The Gyrostar ENV-011D, built by Murata, was chosen. This piezoelectric vibrating prism gyrometer measures angular velocity in a range of ±90 deg/sec - see [10] for details. Orientation must be calculated by integration of the angular velocity over time. The integration leads error accumulation, causing drift over long periods of time. To overcome this problem, external references have to be used, such as the fluxgate compass included in our system but not described here - see [13] for details.

To measure tilt about the x- and y-axis a dual axis AccuStar electronic clinometer, built by Lucas Sensing Systems, was used. The device has a linear range of ± 20º. The dome shaped sensor has a dielectric fluid and an air bubble inside that determine the capacitance in each quadrant, enabling tilt sensing on both axis.

B. The Processing Hardware

To handle the inertial data acquisition, and also enable some processing, a micro-controller based card was built. This card has analog filters, an A/D converter as is based on *Intel*'s 80C196KC micro-controller. The robot's master processing unit has an EISA bus interface, where the card is connected along with another for image acquisition and processing. The *Transtech Parallel Systems* TDMB412 card with a framegrabber module and two *Texas Instruments* TMS320C40 DSPs handles the video processing.

Fig. 3 shows the architecture of the system and the computer that supervises the active vision, moving platform and inertial system. The inertial sensors were mounted inside an acrylic cube, enabling the correct alignment of the gyros, clinometer (mounted on the outside) and accelerometer, as can be seen in the close up Fig. 2. This cube is connected to and continuously monitored by the micro-controller card in the host computer.

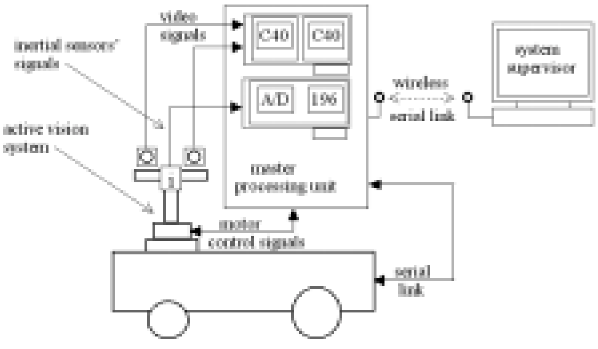

Fig. 3. System Architecture. The inertial system processing board uses the *Master* processing unit as *host computer*.

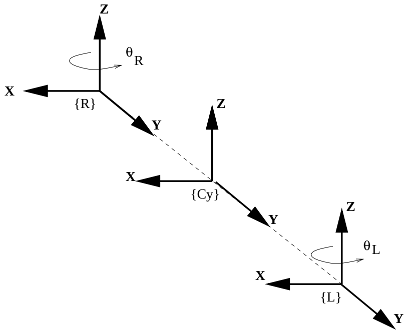

Fig. 4. System Geometry

IV. SYSTEM GEOMETRY

The inertial unit is placed at the middle of the stereo camera baseline, as seen in Fig. 4. The head coordinate frame referential, or Cyclop {Cy}is defined as having the origin at the center of the baseline of the stereo cameras. Each camera position has its own referential, {R} and {L} being for the right and left positions. The cameras have vengeance capabilities, thus having {R} and {L} not only translated along the baseline ({Cy} *y-axis*) but also rotated with respect to their *z-axis*. Notice that in our case we have a common baseline and symmetric vengeance, i.e. $\theta_R=-\theta_L$.

To comply with the typical referential convention used for cameras, two additional referentials, $\{C_R\}$ and $\{C_L\}$, are considered -see Fig. 5.

To express a world point $\vec{P}$, given in the camera referentials, on the Cyclop referential {Cy} we have

$$ {}^{C_y}\vec{P} = {}^{C_y}T_{C_R} . {}^{C_R}\vec{P} = {}^{C_y}T_R . {}^{R}T_{C_R} . {}^{C_R}\vec{P} \tag{1} $$

and

$$ {}^{C_y}\vec{P} = {}^{C_y}T_{C_L} . {}^{C_L}\vec{P} = {}^{C_y}T_L . {}^{L}T_{C_L} . {}^{C_L}\vec{P} \tag{2} $$

where

$$ {}^{R}T_{C_R} = {}^{L}T_{C_L} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{3} $$

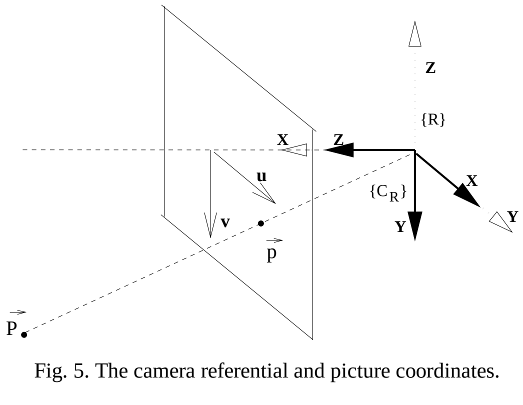

Fig. 5. The camera referential and picture coordinates.

$$^{C_y}T_R = \begin{bmatrix} \cos\theta_R & -\sin\theta_R & 0 & 0 \\ \sin\theta_R & \cos\theta_R & 0 & -\frac{b}{2} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{4}$$

$$^{C_y}T_L = \begin{bmatrix} \cos\theta_L & -\sin\theta_L & 0 & 0 \\ \sin\theta_L & \cos\theta_L & 0 & \frac{b}{2} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5}$$

and b is the baseline distance.

A projection point $\vec{p}=(u,v)$ in each camera image is related with a 3D point $\vec{P}=(X,Y,Z)$ by the perspective relations

$$u = S_u f \frac{X}{Z} \quad \text{and} \quad v = S_v f \frac{Y}{Z}, \tag{6}$$

where u and v are the pixel coordinates with origin at the image center, f is the camera focal distance, S_u and S_v are the scale factors and $\vec{P}$ is expressed in the camera referential.

If we know $\vec{P}=(X,Y,Z)$, finding the projection (u,v) is trivial. The reverse problem involves matching points between the left and right images. Establishing this correspondence will give us enough equations to determine the 3D coordinates, if a few vision system parameters are known. However if visual fixation is used, the geometry is simplified and the reconstruction of the 3D fixated point is simplified, as can be seen in Fig. 6. Notice that the visual fixation can be achieved by controlling the active vision system and the geometry generated by the process allows a fast and robust 3D reconstruction of the fixation point - see [14] and [15] for details.

V. IDENTIFYING THE GROUND PLANE

The clinometer data can be used to determine the orientation of the ground plane. In order to locate this plane in space at least one point belonging to the ground plane must be known.

A. Clinometer gives the ground plane Orientation

When the vehicle is stationary or subject to constant speed the clinometer gives the direction of $\vec{g}$ relative to the Cyclop referential {Cy}. Assuming the ground is leveled, and with α_x and α_y being the sensed angles along the *x* and *y axis*, the normal to the ground plane will be

$$\hat{n} = -\frac{\vec{g}}{\|\vec{g}\|} = \begin{bmatrix} -\sin\alpha_y \\ \sin\alpha_x \\ \sqrt{1-\sin^2\alpha_x - \sin^2\alpha_y} \end{bmatrix}, \tag{7}$$

given in the Cyclop frame of reference. Using this inertial information the equation for the ground plane will be given by

$$\hat{n}.\vec{P} + h = 0, \tag{8}$$

where $\vec{P}$ is a point in the plane and h is the distance from the origin of {Cy}down to the ground plane.

B. Fixing a point in the image gives plane

To obtain a point belonging to the ground plane it is necessary to establish a mechanism to achieve visual fixation. This mechanism was developed in our laboratory and is described in [14] and [15]. If the active vision system fixates in a point that belongs to the ground plane, the ground plane could be determined in the Cyclop referential {Cy} using the clinometer data. Hence, any other correspondent point in the image can be identified as belonging or not to the ground plane.

If the fixation point is identified as belonging to the ground plane, the value of *h* in (8), can be determined. As seen in Fig. 6 (where only α_y is non null to keep the diagram simple) *h* will be given by

$$h = -\hat{n}.\vec{P}_f. \tag{9}$$

C. Segmentation of ground plane

An algorithm for the segmentation of the ground plane can now be presented, based on the solution of (9). Starting with a point of interest in one of the images, say the right image (u_R, v_R), from (6) and (2) the point expressed in the Cyclop referential is given by

$$^{C_y}\vec{P} = {}^{C_y}T_{C_R} . {}^{C_R}\vec{P}(u_R, v_R, \lambda_R), \tag{10}$$

where λ_R represents an unknown value (depending on depth from camera). Substituting (10) in (8), λ_R can be determined and hence $\vec{P}$ is completely known in the Cyclop referential. Expressing $\vec{P}$ in the {CL} referential by

$$^{C_L}\vec{P} = {}^{C_L}T_{C_y} . {}^{C_y}\vec{P} = \left({}^{C_y}T_{C_R}\right)^{-1} . {}^{C_y}\vec{P}, \tag{11}$$

the correspondent point of interest (u_L, v_L) generated by the projection of $\vec{P}$ in the left image is given by

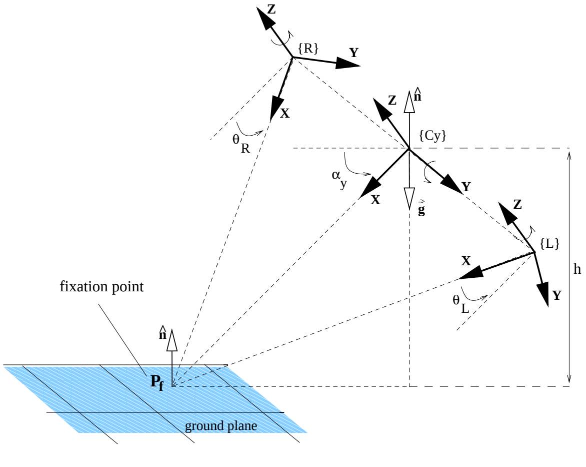

Fig. 6. Ground plane point fixated. The point $\vec{P}_f$ in the ground plane is visualized by the active vision system. The geometry of this visualization corresponds to a state, named visual fixation.

$$u_L = S_u f \frac{X}{Z} \quad \text{and} \quad v_L = S_v f \frac{Y}{Z}. \tag{12}$$

The correspondent point and its neighborhood in the left image can then be tested for a match with the original point of interest in the right image. If there is a match, the point belongs to the ground plane. If there is no match the point must be something other than the floor, possibly an obstacle.

D. Ground point computation

When visual fixation is obtained for a ground point and assuming symmetric vengeance (i.e. $\theta=\theta_R=-\theta_L$.) from (9) and (7) we have

$$h = -\hat{n}.\vec{P}_f = \frac{b \sin\alpha_y \cos\theta}{2\sin\theta} \tag{13}$$

as can easily be seen in Fig. 6 (where α_x is null, but (13) still holds for any α_x).

This value of h will be used to determine if other points in the image belong or not to the level plane passing the image center point (i.e. fixation point).

Taking λ_R out of (13) and (10) and substituting in (11) and (12) we get

$$u_L = \frac{-fS_u\left(k_d u_R + k_e v_R\right)}{k_a u_R + k_b v_R + k_c} \tag{14}$$

and

$$u_L = \frac{k_c v_R}{k_a u_R + k_b v_R + k_c} \tag{15}$$

where

$$k_a = -2S_v \sin\theta \sin\alpha_y\,, \tag{16}$$

$$k_b = 2S_u \sin^2\theta \sqrt{1-\sin^2\alpha_x - \sin^2\alpha_y}\,, \tag{17}$$

$$k_c = fS_u S_v\left(\cos\theta\sin\alpha_y - \sin\theta\sin\alpha_x\right), \tag{18}$$

$$k_d = -S_v\left(\cos\theta\sin\alpha_y + \sin\theta\sin\alpha_x\right), \tag{19}$$

$$k_e = 2S_u \cos\theta\sin\theta \sqrt{1-\sin^2\alpha_x - \sin^2\alpha_y}\,. \tag{20}$$

This set of equations allows a quick computation of the algorithm for a given a set of stereo images.

E. Results

Fig. 8 shows a pair of stereo images where fixation was obtained for a ground plane point.

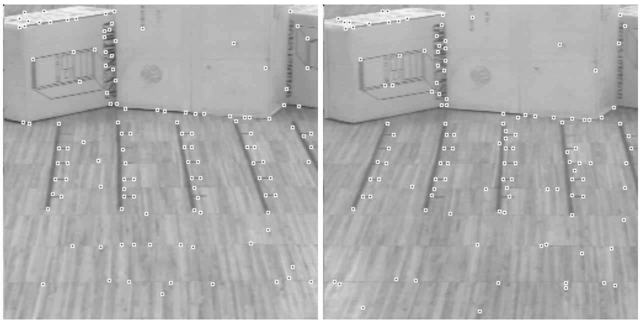

Fig. 7. Initial points.

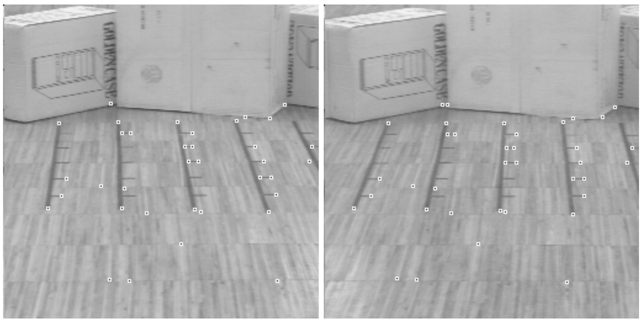

Fig. 8. Example of ground point detection.

In this example α_x is null and α_y=14.0º, θ=1.60º and b=29.6cm. Making $h=b\times\cot(\theta)\times\sin(\alpha_y)/2\cong 129$cm. The ground plane is thus determined.

The points of interest in the right image can be parsed as described in the previous section. Fig. 8 shows the matched ground plane points of interest. The initial points in each image shown in Fig. 7 were obtained using SUSAN corner detector [16]. The results are still preliminary, and more work in needed to improve the algorithm's performance, allowing a better ground plane segmentation.

VI. CONCLUSIONS

The integration of artificial inertial systems in autonomous systems opens a new field for the development of applications based or related with inertial information. Using the information provided by an artificial inertial system, several human vestibular functions could be implemented in an autonomous system or active vision system.

In this article we described a system that integrates inertial sensors and active vision systems used in autonomous vehicles. As a first approach information about the ground plane extracted from the inertial sensors is used to identify the floor plane using the simulation of visual fixation with an active vision system. Fixating the visual system on a point in the ground plane, the 3D position of the plane in space is obtained. Any other correspondent point in the stereo image can be identified as belonging or not to the ground plane. Segmentation of the image is therefore accomplished. Some preliminary results were presented.

VII. ACKNOWLEDGMENTS

Financial support for this work was partially provided by JNICT-Junta Nacional de Investigação Científica e Tecnológica, on the framework of the SIRMA project and as an M.Sc. student grant (BM/6531/95) given by PRAXIS XXI, JNICT, to J. Lobo.

VIII. REFERENCES

[1] H. Carpenter, *Movements of the Eyes*, London Pion Limited, ISBN 0 85086 109 8, 2nd edition.

[2] W.T. Russel, "Theory of inertial navigation", in *Inertial Guidance*, ed. George R. Pitman, John Wiley

& Sons, 1962, pp. 16-46.

[3] T. Viéville and O.D. Faugeras, “Computation of Inertial Information on a Robot”, Fifth International Symposium on Robotics Research, Hirofumi Miura and Suguru Arimoto editor, MIT-Press, 1989, pp. 57-65.

[4] S. Cooper and H. Durrant-Whyte, "A Kalman Filter Model for GPS Navigation of Land Vehicles", in *Proceedings of the 1994 IEEE International Conference on Intelligent Robots and Systems*, pp. 157-163.

[5] Y. Tonouchi, T. Tsubouchi and S. Arimoto, "Fusion of Dead-reckoned Positions with Workspace Model for a Mobile Robot by Bayesian Inference", in *Proceedings of the 1994 IEEE International Conference on Intelligent Robots and Systems*, pp. 1347-1354.

[6] IC Sensors, "Silicon Accelerometers", Technical Note TN-008.

[7] Summit Instruments, "34100A Theory of Operation", Technical Note 402.

[8] Analog Devices, "ADXL05", data sheet.

[9] R.P.G. Collinson, *Introduction to Avionics*, Chapman & Hall, ISBN 0 412 48250 9, 1996, pp.187-204.

[10] Murata MFG. Co. Ltd, "Piezoelectric Vibrating Gyroscope GYROSTAR", Cat. No. S34E, 1991

[11] B. Barshan, "Orientation Estimate for Mobile Robots using Gyroscopic Information", in *Proceedings of the 1994 IEEE International Conference on Intelligent Robots and Systems*, pp. 1867-1874.

[12] B. Barshan and H. Durrant-Whyte, "Inertial Navigation Systems for Mobile Robots”, *IEEE Trans. Robotics and Automation*, vol. 11, no. 3, June 1995, pp. 328-342.

[13] T.J. Peters, "Automobile Navigation Using a Magnetic Flux-Gate Compass", *IEEE Trans. Vehicular Technology*, vol. VT-35, no. 2, May 1986, pp. 41-47.

[14] J. Dias, C. Paredes, I. Fonseca *et al*. "Simulating Pursuit with Machines", In *Proceedings of the 1995 IEEE Conference on Robotics and Automation*, Japan, pp. 472-477.

[15] C. Paredes, J. Dias, A. de Almeida. "Detecting Movements Using Fixation", in *Proceedings of the 2nd Portuguese Conference on Automation and Control*, Oporto, September 1996, pp. 741-746.

[16] S.M. Smith and J.M. Brady. "SUSAN - a new approach to low level image processing", in *Int. Journal of Computer Vision*, 1997. In publication.